

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-296891  
 (43)Date of publication of application : 26.10.2001

(51)Int.Cl. G10L 15/28  
 G10L 15/00  
 G10L 15/20  
 G10L 21/02

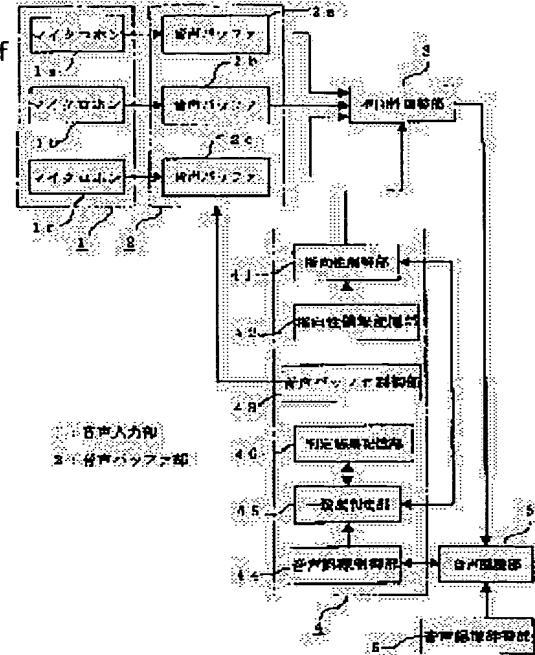
(21)Application number : 2000-112942 (71)Applicant : MITSUBISHI ELECTRIC CORP  
 (22)Date of filing : 14.04.2000 (72)Inventor : NOGI KAZUYUKI

## (54) METHOD AND DEVICE FOR VOICE RECOGNITION

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To cope with the uttering of speakers located in plural directions without utilizing a switch for speaker discrimination.

**SOLUTION:** Keyword input signals from a voice input section 1 are stored in a voice buffer section 2. A directivity adjusting section 3 changes the directivity of the signals and a voice recognition section 5 conducts recognition processes for the signals. A degree of agreement discrimination section 45 discriminates a directivity having high degree of agreement among the recognition results. By setting the directivity, excellent voice recognition is attained.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

**BEST AVAILABLE COPY**

[Date of extinction of right]

Copyright (C): 1998,2003 Japan Patent Office

(19) 日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-296891

(P2001-296891A)

(43) 公開日 平成13年10月26日 (2001.10.26)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード <sup>8</sup> (参考)
G 10 L 15/28		C 10 L 3/00	511 5D015
15/00			551J
15/20			571J
21/02			571M
		3/02	301B

審査請求 未請求 請求項の数 4 O L (全 14 頁)

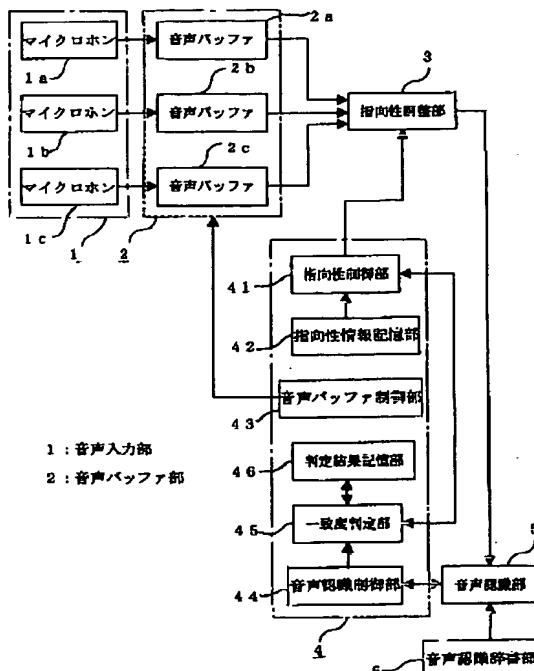
(21) 出願番号	特願2000-112942 (P2000-112942)	(71) 出願人	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目2番3号
(22) 出願日	平成12年4月14日 (2000.4.14)	(72) 発明者	野木 和行 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社内
		(74) 代理人	100093562 弁理士 久玉 俊英
			F ターム(参考) 5D015 AA04 DD02 EE01 KK01 LL09 LL10

### (54) 【発明の名称】 音声認識方法および装置

#### (57) 【要約】

【課題】 複数の方向に存在する話者の発声にも、話者判別用のスイッチなしで対応できるようにする。

【解決手段】 音声入力部1からのキーワード入力信号を音声バッファ部2に格納し、その音声信号に対して指向性調整部3で指向性を変え、音声認識部5で認識処理する。その認識結果の中から、一致度判定部4で最も一致度の高い指向性を判定する。この指向性に設定することにより、良好な音声認識を行うことができる。



【特許請求の範囲】

【請求項1】 音声入力部と、この音声入力部からの信号を保存する音声バッファ部と、この音声バッファ部に保存された信号から指向性を変更した信号を生成する指向性調整部と、この指向性調整部により指向性を変更した信号に対して音声認識処理を行う音声認識部と、上記音声バッファ部に保存された信号を再生する音声バッファ制御部と、上記指向性調整部にて変更する複数の指向性情報を記憶した指向性情報記憶部と、この指向性情報記憶部に記憶された指向性情報から1つを選択して上記指向性調整部を制御する指向性制御部と、上記音声認識部における認識対象を記憶した音声認識辞書部と、上記音声認識部を制御して認識処理を複数回実行させる音声認識制御部と、変更した複数の指向性における認識結果の一一致度を記憶する判定結果記憶部と、この判定結果記憶部に記憶された認識結果の中から一致度が最も高い指向性を判定する一致度判定部とを備えた音声認識装置。

【請求項2】 音声入力部と、この音声入力部からの信号を保存する音声バッファ部と、この音声バッファ部に保存された信号から指向性および利得を変更した信号を生成する指向性利得調整部と、この指向性利得調整部により指向性および利得を変更した信号に対して音声認識処理を行う音声認識部と、上記音声バッファ部に保存された信号を再生する音声バッファ制御部と、上記指向性利得調整部にて変更する複数の指向性情報および利得情報を記憶した指向性利得情報記憶部と、この指向性利得情報記憶部に記憶された指向性情報および利得情報から各1つを選択して上記指向性利得調整部を制御する指向性利得制御部と、上記音声認識部における認識対象を記憶した音声認識辞書部と、上記音声認識部を制御して認識処理を複数回実行させる音声認識制御部と、変更した複数の指向性および利得における認識結果の一一致度を記憶する判定結果記憶部と、この判定結果記憶部に記憶された認識結果の中から一致度が最も高い指向性および利得を判定する一致度判定部とを備えた音声認識装置。

【請求項3】 入力された音声信号に対して認識処理を行い、あるキーワードが認識されたかどうかを判定するステップ、キーワードが認識されたとき、そのキーワードの音声信号に対して指向性を変更した複数の認識処理を行い、それぞの認識結果の一一致度を得るステップ、これらの認識結果の中から一致度が最も高い指向性を判定するステップ、およびこの判定した指向性に設定してその後の音声認識を行うステップを有する音声認識方法。

【請求項4】 入力された音声信号に対して認識処理を行い、あるキーワードが認識されたかどうかを判定するステップ、キーワードが認識されたとき、そのキーワードの音声信号に対して指向性および利得を変更した複数の認識処理を行い、それぞの認識結果の一一致度を得るステップ、これらの認識結果の中から一致度が最も高い

指向性および利得を判定するステップ、およびこの判定した指向性および利得に設定してその後の音声認識を行うステップを有する音声認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声で制御される制御対象機器に与える入力指示の語彙を認識する、車載用などの音声認識装置に関するものである。

【0002】

【従来の技術】図8は、従来の車載用音声認識装置のシステム構成を示すブロック図である。以下、図に従って説明する。車両等の運転中にナビゲーション機器やオーディオ機器等を操作するに際して、スイッチ操作による運転者の負担を軽減するために、運転者など話者（発話者）の発声した音声を認識して、接続された機器に入力指示する音声認識装置がある。

【0003】1は話者の発声した音声を電気信号に変換する音声入力部で、無指向性の複数のマイクロホン1a～1cからなっている。3はマイクロホン1a～1cで検出された信号を調整して音声入力部1の指向性を話者の方向に調整した音声信号を出力するための指向性調整部である。5はナビゲーション機器やオーディオ機器等の入力部に接続された音声認識部である。8は話者の方向を検出する方向検出手段で、ルームミラーの角度や座席スライド位置、リクライニング角度などから音声の方向を検出する。41は方向検出手段8の検出結果に基いて指向性調整部3を制御する指向性制御部である。

【0004】次に、動作について説明する。図9は、従来の車載用音声認識装置の動作内容を示すフローチャートである。最初にステップS0において、音声認識開始の操作が行われる。次にステップS1において、方向検出手段8で話者の方向を検出し、話者の位置情報を取得する。次にステップS2において、ステップS1にて取得した話者の位置情報を基に指向性制御部41が指向性を話者の方向に変更するように指向性調整部3を制御する。次にステップS3において、マイクロホン1a～1cから話者の音声を入力する。続いてステップS4において、指向性調整部3が指向性を変更した音声に対して音声認識部5が認識処理を実行する。そしてステップS5において、音声認識部5から認識結果を出力する。

【0005】

【発明が解決しようとする課題】従来の音声認識装置では、指向性を調整する場合、話者の位置を特定するために、話者の方向を検出する手段が必要となる。この方向検出手段として従来は、車載用の場合、ルームミラーの角度や座席スライド位置、リクライニング角度などから検出した。そのため、話者が運転者であることに限定されてしまい、運転者以外が話者となる場合は上記方向検出手段では検出できない。運転者以外をも話者としてその音声を検出する場合、方向検出手段の構成が複雑とな

るだけでなく、誰が発話するのかを検出しなければならないのでその手段として話者判別用のスイッチなどが必要となる。また、これらの構成を実現した場合においても、指定した指向性が必ずしも音声認識処理において高い性能が得られる特性であるとは限らない。

【0006】本発明は、話者の音声以外のノイズを低減して音声認識性能を向上させると共に、複数の方向に存在する話者の発声にも、話者判別用のスイッチなしで対応できる利便性の高い音声認識方法、装置を提供することを目的とする。

【0007】

【課題を解決するための手段】請求項1に係る音声認識装置は、音声入力部と、音声入力部からの信号を保存する音声バッファ部と、音声バッファ部に保存された信号から指向性を変更した信号を生成する指向性調整部と、指向性調整部により指向性を変更した信号に対して音声認識処理を行う音声認識部と、音声バッファ部に保存された信号を再生する音声バッファ制御部と、指向性調整部にて変更する複数の指向性情報を記憶した指向性情報記憶部と、指向性情報記憶部に記憶された指向性情報から1つを選択して指向性調整部を制御する指向性制御部と、音声認識部における認識対象を記憶した音声認識辞書部と、音声認識部を制御して認識処理を複数回実行させる音声認識制御部と、変更した複数の指向性における認識結果の一致度を記憶する判定結果記憶部と、判定結果記憶部に記憶された認識結果の中から一致度が最も高い指向性を判定する一致度判定部とを備えたものである。

【0008】請求項2に係る音声認識装置は、音声入力部と、音声入力部からの信号を保存する音声バッファ部と、音声バッファ部に保存された信号から指向性および利得を変更した信号を生成する指向性利得調整部と、指向性利得調整部により指向性および利得を変更した信号に対して音声認識処理を行う音声認識部と、音声バッファ部に保存された信号を再生する音声バッファ制御部と、指向性利得調整部にて変更する複数の指向性情報および利得情報を記憶した指向性利得情報記憶部と、指向性利得情報記憶部に記憶された指向性情報および利得情報から各1つを選択して指向性利得調整部を制御する指向性利得制御部と、音声認識部における認識対象を記憶した音声認識辞書部と、音声認識部を制御して認識処理を複数回実行させる音声認識制御部と、変更した複数の指向性および利得における認識結果の一致度を記憶する判定結果記憶部と、判定結果記憶部に記憶された認識結果の中から一致度が最も高い指向性および利得を判定する一致度判定部とを備えたものである。

【0009】請求項3に係る音声認識方法は、入力された音声信号に対して認識処理を行い、あるキーワードが認識されたかどうかを判定するステップ、キーワードが認識されたとき、そのキーワードの音声信号に対して指

向性を変更した複数の認識処理を行い、それぞれの認識結果の一一致度を得るステップ、これらの認識結果の中から一致度が最も高い指向性を判定するステップ、およびこの判定した指向性に設定してその後の音声認識を行うステップを有するものである。

【0010】請求項4に係る音声認識方法は、入力された音声信号に対して認識処理を行い、あるキーワードが認識されたかどうかを判定するステップ、キーワードが認識されたとき、そのキーワードの音声信号に対して指向性および利得を変更した複数の認識処理を行い、それぞれの認識結果の一一致度を得るステップ、これらの認識結果の中から一致度が最も高い指向性および利得を判定するステップ、およびこの判定した指向性および利得に設定してその後の音声認識を行うステップを有するものである。

【0011】

【発明の実施の形態】実施の形態1. 以下、この発明の実施の形態を、車載用の音声認識装置について説明する。図1は、この発明の実施の形態1における音声認識装置のシステム構成を示すブロック図である。図において、1は話者が発した音声を電気信号（以下、音声信号と呼ぶ）に変換する音声入力部で、無指向性の複数（ここでは3つ）のマイクロホン1a～1cからなる。図2は、車両へのマイクロホンの取付位置の例を示す平面図である。車両1内で、3つのマイクロホン1a～1cをダッシュボード12上中央部へ等間隔に三角形をなすように設置する。例えば、マイクロホン1a～1cの各出力信号のゲインバランスを調整することにより、指向性を真正面や運転席13方向、あるいは助手席14方向に変更することが可能となる。

【0012】図1へ戻り、2は音声入力部1からの音声信号を保存する音声バッファ部で、それぞれマイクロホン1a～1cに対応して設けられた複数の音声バッファ2a～2cからなる。3は音声バッファ部2に保存された音声信号を調整して、指向性を変更した音声信号を出力する指向性調整部、5は指向性調整部3で指向性を変更された音声信号に対して音声認識処理を実行する音声認識部であり、認識結果と次に述べる音声認識辞書部6に保存されたデータとの一致度を出力する。6は認識対象を記憶する音声認識辞書部であり、音声認識部5の音声認識処理における基準となるデータが保存されている。

【0013】4は音声バッファ部2と指向性調整部3と音声認識部5を制御する制御部、43は音声バッファ部2での音声信号の保存と再生を制御する音声バッファ制御部、41は指向性調整部3での指向性変更を制御する指向性制御部、42は指向性制御部41による指向性制御のための複数の指向性情報を記憶する指向性情報記憶部であり、例えば正面を0°とし、5°間隔で±90°までを記憶しておき、指向性制御部41がその中から1

つづつ選択して制御を行う。4 4は音声認識部5の認識処理の開始や中止、および認識結果と一致度の取得を行う音声認識制御部、4 5は指向性制御部4 1からの指向性情報および音声認識制御部4 4からの認識結果と一致度から、どの指向性が最適であるかを判定する一致度判定部、4 6は音声認識部での認識結果と一致度、および一致度判定部4 5での判定結果を記憶する判定結果記憶部であり、上記4 1～4 6で制御部4を構成している。

【0014】次に、動作について説明する。図3は、図1に示した音声認識装置の動作内容を示すフローチャートである。最初にステップA 0において、制御部4の各部の初期化および処理の開始操作が実行される。次にステップA 1において、指向性制御部4 1からの制御により指向性調整部3の指向性設定を無指向性に設定する。次にステップA 2において、マイクロホン1 a～1 cに入力されて音声信号に変換されたそれぞれの信号を音声バッファ制御部4 3からの制御により音声バッファ2 a～2 cに格納し、この格納された音声信号を音声バッファ制御部4 3からの制御により再生し、この音声信号を、無指向性に設定された指向性調整部3に入力し、指向性調整部3の出力を音声認識部5に入力する。音声認識部5では入力された音声信号に対して、音声認識制御部4 4からの制御により、音声操作の開始コマンドとなるキーワード、例えば「認識スタート」を認識する処理を実行する。

【0015】次にステップA 3において、音声認識部5の認識結果に基づき、音声認識制御部4 4は、キーワード「認識スタート」が認識されたのかを判定し、認識されなかった場合はステップA 2に戻り、再度音声入力処理およびキーワード認識処理を実行する。認識された場合はステップA 4に進む。次に、ステップA 4へ進んだときは、音声バッファ制御部4 3からの制御により、音声バッファ2 a～2 cへの音声入力を停止し、キーワード「認識スタート」が認識された時のその音声信号を格納する。

【0016】次にステップA 5において、指向性制御部4 1からの制御により、指向性調整部3の指向性設定を指向性情報記憶部4 2に記憶された、例えば正面0°方向に設定する。次にステップA 6において、ステップA 4で音声バッファ2 a～2 cに格納されたキーワード「認識スタート」の音声信号を、音声バッファ制御部4 3からの制御により再生し、指向性調整部3にて指向性制御部4 1が設定した指向性をもつ音声信号を生成し、音声認識部5にてキーワード認識処理を再度実行し、音声認識制御部4 4が認識処理の結果と一致度を音声認識部5から取得して、一致度判定部4 5に送信する。一致度判定部4 5は現在設定されている指向性情報と認識結果と一致度を判定結果記憶部4 6へ送信して記憶する。

【0017】次にステップA 7において、指向性情報記

憶部4 2に記憶された全ての指向性についての再認識処理および認識結果と一致度の取得が終了していない場合はステップA 5に戻り、指向性情報記憶部4 2に記憶された全ての指向性について終了するまで繰り返す。全ての指向性について再認識処理および認識結果と一致度の取得が終了した場合はステップA 8に進む。

【0018】次に、ステップA 8へ進んだときは、判定結果記憶部4 6に記憶された全ての指向性についての認識結果と一致度から、一致度判定部4 5は、認識結果が正解、すなわちキーワード「認識スタート」であり、かつ最も一致度の高い指向性はどれであるかを判定し、指向性制御部4 1は、一致度判定部4 5が判定した指向性となるように指向性調整部3を制御する。

【0019】次にステップA 9において、ステップA 4で停止した音声バッファ2 a～2 cへの音声入力を再開する。すなわちマイクロホン1 a～1 cに入力されて音声信号に変換されたそれぞれの信号を音声バッファ制御部4 3からの制御により音声バッファ2 a～2 cに格納し、この格納された音声信号を音声バッファ制御部4 3からの制御により再生し、ステップA 8で認識結果が正解であってかつ最も一致度の高い指向性に設定された指向性調整部3に音声信号を入力し、指向性調整部3の出力を音声認識部5に入力する。音声認識部5では入力された音声信号に対して、音声認識制御部4 4からの制御により音声認識辞書部6に格納された認識語彙を認識する処理を実行する。次にステップA 10において、音声認識部5は認識処理の結果を出力し、図示外のオーディオ機器等の操作を行う。なお、ステップA 8で、図示外のディスプレイまたはランプにより、キーワード「認識スタート」の認識完了と指向性の設定方向を表示するようにしておけば、話者がその表示を確認して、ステップA 9で、続くコマンドを入力することができる。

【0020】以上のように実施の形態1の音声認識方法、装置においては、キーワード「認識スタート」を認識した時点の音声バッファの音声信号を用いて指向性を変更し、音声認識における一致度から話者の方向を判定して、話者の音声を抽出するため、話者の方向が定まっていない場合においても話者の音声を有効に抽出し、認識する事が可能である。また、話者の方向検出手段が不要であり、話者判別用のスイッチあるいは方向検出用のセンサなどのコスト削減が可能となる。

【0021】実施の形態2、図4は、この発明の実施の形態2における音声認識装置のシステム構成を示すブロック図である。本実施の形態では、実施の形態1で行った指向性を変化させ一致度の最大のものを選ぶ方法に加えて、利得変化すなわち信号レベルを変化させて一致度を見る方法を用いている。図4では図1の指向性調整部3、指向性制御部4 1および指向性記憶部4 2に代えて、それぞれ指向性利得調整部3 1、指向性利得制御部4 7および指向性利得記憶部4 8を設けている。

【0022】図4において、31は音声バッファ部2に保有された音声信号を調整して指向性および利得を変更した音声信号を出力する指向性利得調整部、47は指向性利得調整部31での指向性と利得の変更を制御する指向性利得制御部、48は指向性利得制御部47の指向性と利得の制御において複数の指向性情報と複数の利得情報を記憶する指向性利得情報記憶部であり、例えば正面を0°とし、5°間隔で±90までを記憶するとともに、初期利得を0dBとし、3dB間隔で±15dBまでを記憶している。

【0023】音声認識部5は、指向性利得調整部31で指向性と利得を変更された音声信号に対して音声認識処理を実行する。一致度判定部45は、指向性利得制御部47からの指向性情報と利得情報および音声認識制御部44からの認識結果と一致度から、どの指向性および利得が最適であるかを判定する。制御部4は、43～48で構成されている。その他の部分は図1と同様であるので説明を省略する。

【0024】次に、動作について説明する。図5は、図4に示した音声認識装置の動作内容を示すフローチャートである。最初にステップB0において、制御部4の各部の初期化および処理の開始操作が実行される。次にステップB1において、指向性利得制御部47からの制御により指向性利得調整部31の指向性および利得設定を無指向性および初期利得に設定する。次にステップB2において、マイクロホン1a～1cに入力されて音声信号に変換されたそれぞれの信号を音声バッファ制御部43からの制御により音声バッファ2a～2cに格納し、この格納された音声信号を音声バッファ制御部43からの制御により再生し、この音声信号を、無指向性および初期利得に設定された指向性利得調整部31に入力し、指向性利得調整部31の出力を音声認識部5に入力する。音声認識部5では入力された音声信号に対して、音声認識制御部44からの制御により、音声操作の開始コマンドとなるキーワード、例えば「認識スタート」を認識する処理を実行する。

【0025】次にステップB3において、音声認識部5の認識結果に基づき、音声認識制御部44は、キーワード「認識スタート」が認識されたのかを判定し、認識されなかった場合はステップB2に戻り再度音声入力処理およびキーワード認識処理を実行する。認識された場合はステップB4に進む。次に、ステップB4へ進んだときは、音声バッファ制御部43からの制御により、音声バッファ2a～2cへの音声入力を停止し、キーワード「認識スタート」が認識された時のその音声信号を格納する。

【0026】次にステップB5において、指向性利得制御部47からの制御により、指向性利得調整部31の指向性設定を指向性利得情報記憶部48に記憶された例えば正面0°方向に設定する。次にステップB6におい

て、ステップB4で音声バッファ2a～2cに格納されたキーワード「認識スタート」の音声信号を音声バッファ制御部43からの制御により再生し、指向性利得調整部31にて指向性利得制御部47が設定した指向性をもつ音声信号を生成し、音声認識部5にてキーワード認識処理を再度実行し、音声認識制御部44が認識処理の結果と一致度を音声認識部5から取得して、一致度判定部45に送信する。一致度判定部45は現在設定されている指向性情報と認識結果と一致度を判定結果記憶部46へ送信して記憶させる。

【0027】次にステップB7において、指向性利得情報記憶部48に記憶された全ての指向性についての再認識処理および認識結果と一致度の取得が終了していない場合はステップB5に戻り、指向性利得情報記憶部48に記憶された全ての指向性について再認識処理および認識結果と一致度の取得が終了するまで繰り返す。全ての指向性について再認識処理および認識結果と一致度の取得が終了した場合はステップB8に進む。

【0028】次に、ステップB8へ進んだときは、判定結果記憶部46に記憶された全ての指向性についての認識結果と一致度から、一致度判定部45は、認識結果が正解、すなわちキーワード「認識スタート」であり、かつ最も一致度の高い指向性はどれであるかを判定し、指向性利得制御部47は、一致度判定部45が判定した指向性となるように指向性利得調整部31を制御する。

【0029】次にステップB9において、指向性利得制御部47からの制御により、指向性利得調整部31の利得設定を指向性利得情報記憶部48に記憶された、例えば初期利得より3dB高い利得に設定する。この場合の利得調整は、ステップB8にて判定された指向性の方向についてのみ利得が調整されるものとする。

【0030】次にステップB10において、ステップB4で音声バッファ2a～2cに格納されたキーワード「認識スタート」の音声信号を音声バッファ制御部43からの制御により再生し、指向性利得調整部31にて指向性利得制御部47が設定した指向性および利得をもつ音声信号を生成し、音声認識部5にてキーワード認識処理を再度実行し、音声認識制御部44が認識処理の結果と一致度を音声認識部5から取得し、一致度判定部45に送信する。一致度判定部45は現在設定されている指向性情報と利得情報と認識結果と一致度を判定結果記憶部46へ送信して記憶させる。

【0031】次にステップB11において、指向性利得情報記憶部48に記憶された全ての利得についての再認識処理および認識結果と一致度の取得が終了していない場合はステップB9に戻り指向性利得情報記憶部48に記憶された全ての利得に対する再認識処理および認識結果と一致度の取得が終了するまで繰り返す。全ての利得について再認識処理および認識結果と一致度の取得が終了した場合はステップB12に進む。

【0032】次に、ステップB12へ進んだときは、判定結果記憶部46に記憶された全ての利得についての認識結果と一致度から、一致度判定部45は、認識結果が正解であり、かつ最も一致度の高い利得はどれであるかを判定し、指向性利得制御部47は、一致度判定部45が判定した指向性および利得となるように指向性利得調整部31を制御する。

【0033】次にステップB13において、ステップB4で停止した音声バッファ2a～2cへの音声入力を再開する。すなわちマイクロホン1a～1cに入力されて音声信号に変換されたそれぞれの信号を音声バッファ制御部43からの制御により音声バッファ2a～2cに格納し、この格納された音声信号を音声バッファ制御部43からの制御により再生し、ステップB12で認識結果が正解であってかつ最も一致度の高い指向性および利得に設定された指向性利得調整部31に音声信号を入力し、指向性利得調整部31の出力を音声認識部5に入力する。音声認識部5では入力された音声信号に対して、音声認識制御部44からの制御により音声認識辞書部6に格納された認識語彙を認識する処理を実行する。次にステップB14において、音声認識部5は認識処理の結果を出力し、図示外のオーディオ機器等の制御を行う。

【0034】以上のように実施の形態2の音声認識方法、装置においては、音声認識における一致度を用いて話者の方向を判定し、更に音声認識における一致度を用いて最適な利得、すなわち音声認識に最適な入力信号レベルを判定し、話者の音声を抽出するため、話者の方向が定まっていない場合においても話者の音声を有効に抽出するとともに、話者からマイクロホンまでの距離などのためにマイクロホンへの音声入力レベルが異なる場合でも、音声認識に最適な入力信号レベルで認識処理を実行する事が可能である。また、話者の方向検出手段が不要であり、方向検出手用のセンサなどのコスト削減が可能となる。

【0035】実施の形態3、図6は、この発明の実施の形態3における音声認識装置のシステム構成を示すブロック図である。本実施の形態は、音声でナビゲーション装置を制御する例を示す。図6では、図4に示したものに加えてナビゲーション装置7を示している。図6において、7は音声認識の結果に基づいて制御部4により、種々の操作が実行されるナビゲーション装置である。その他は図4と同様であるので説明を省略する。

【0036】次に、動作について説明する。図7は、図6に示した音声認識装置の動作内容を示すフローチャートである。ステップC0～C12は、図5のステップB0～B12と同様であるので説明を省略する。ただし、図5ではキーワードの例を「認識スタート」として説明したが、図7では別のキーワード、例えば「ナビゲーション」を用いる。

【0037】ステップC12に続くステップC13にお

いて、ステップC4で停止した音声バッファ2a～2cへの音声入力を再開する。すなわちマイクロホン1a～1cに入力されて音声信号に変換されたそれぞれの信号を音声バッファ制御部43からの制御により音声バッファ2a～2cに格納し、この格納された音声信号を音声バッファ制御部43からの制御により再生し、ステップC12で認識結果が正解であり、かつ最も一致度の高い指向性および利得に設定された指向性利得調整部31に音声信号を入力し、指向性利得調整部31の出力を音声認識部5に入力する。音声認識部5では入力された音声信号に対して、音声認識制御部44からの制御により、音声認識辞書部6に格納されたナビゲーション装置7の制御コマンド語彙、例えば「詳細表示」、「広域表示」、「目的地設定」などを認識する処理を実行する。

【0038】次にステップC14において、音声認識制御部44が音声認識部5の認識結果、例えば「詳細表示」を取得し、認識結果「詳細表示」に対応した制御信号をナビゲーション装置7に送信する。次にステップC15において、ナビゲーション装置7が受信した認識結果「詳細表示」に対応した制御信号に応じて表示画面を詳細表示する処理を実行する。

【0039】以上のように実施の形態3の音声認識方法、装置においては、キーワード「ナビゲーション」を認識した時点の音声バッファの音声信号を用い指向性を変更して音声認識における一致度の最も高い指向性を判定し、これを話者の方向とし、さらに一致度の最も高い利得を判定して以後の音声認識処理を実行するため、音声認識の開始スイッチなどは不要であり、操作が簡便となる。また、キーワード認識処理における一致度の最も高い指向性および利得を判定するため、話者の方向が運転者に限定されず、助手席からの音声操作も可能となり、以後のコマンド認識処理において最適な指向性と利得で認識処理を実行可能であり、認識性能の向上が可能である。また、話者の方向検出手段が不要であり、方向検出手用のセンサなどのコスト削減が可能となる。

【0040】

【発明の効果】請求項1に係る音声認識装置によれば、指向性を変更した信号に対して音声認識処理を行う音声認識部と、認識結果の中から一致度が最も高い指向性を判定する一致度判定部を備えているので、一致度判定部で一致度が最高と判定した指向性に設定することにより、話者が複数存在してその方向が定まっていない場合でも話者判別用のスイッチやセンサなしで、話者以外のノイズを低減して話者の音声を有効に抽出でき、音声認識性能が高く、利便性の高い音声認識装置が得られる。

【0041】請求項2に係る音声認識装置によれば、指向性および利得を変更した信号に対して音声認識処理を行う音声認識部と、認識結果の中から一致度の最も高い指向性と利得を判定する一致度判定部を備えているので、一致度判定部で一致度が最高と判定した指向性およ

び利得に設定することにより、話者の方向が定まっておらず、また音声入力部（マイクロホン）への入力レベルが大きいあるいは小さい場合でも話者判別用のスイッチやセンサなしで、話者の音声を有効に抽出するとともに適切な信号レベルで認識処理を行うことができ、音声認識性能が高く、利便性の高い音声認識装置が得られる。

【0042】請求項3に係る音声認識方法によれば、認識されたキーワードの音声信号に対して指向性を変更した複数の認識処理を行い、その認識結果の中から一致度が最も高い指向性を判定し、この指向性に設定して以後の音声認識を行うので、話者の方向が定まっていない場合でも話者判別用のスイッチやセンサなしで、話者の音声を有効に抽出でき、音声認識性能が高く、利便性の高い音声認識方法が得られる。

【0043】請求項4に係る音声認識方法によれば、認識されたキーワードの音声信号に対して指向性と利得を変更した複数の認識処理を行い、その認識結果の中から一致度が最も高い指向性および利得を判定し、この指向性および利得に設定して以後の音声認識を行うので、話者の方向が定まっておらず、また音声入力部への入力レベルが大きいあるいは小さい場合でも話者判別用のスイッチやセンサなしで、話者の音声を有効に抽出するとともに適切な信号レベルで認識処理を行うことができ、音声認識性能が高く、利便性の高い音声認識方法が得られる。

【図面の簡単な説明】

【図1】 この発明の実施の形態1における音声認識装置のシステム構成を示すブロック図である。

【図2】 図1の音声認識装置のマイクロホンの取付位置を示す平面図である。

【図3】 図1の音声認識装置の動作内容を示すフローチャートである。

【図4】 この発明の実施の形態2における音声認識装置のシステム構成を示すブロック図である。

【図5】 図4の音声認識装置の動作内容を示すフローチャートである。

【図6】 この発明の実施の形態3における音声認識装置のシステム構成を示すブロック図である。

【図7】 図6の音声認識装置の動作内容を示すフローチャートである。

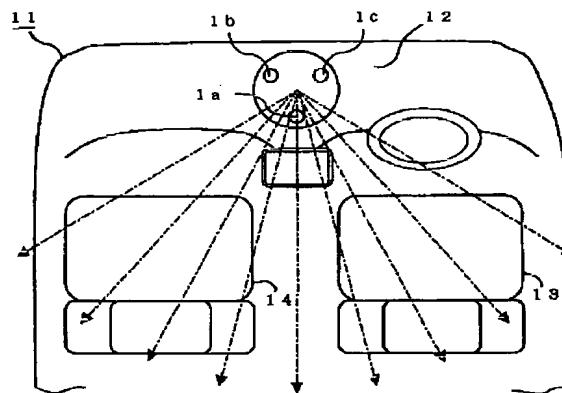
【図8】 従来の音声認識装置のシステム構成を示すブロック図である。

【図9】 図8の音声認識装置の動作内容を示すフローチャートである。

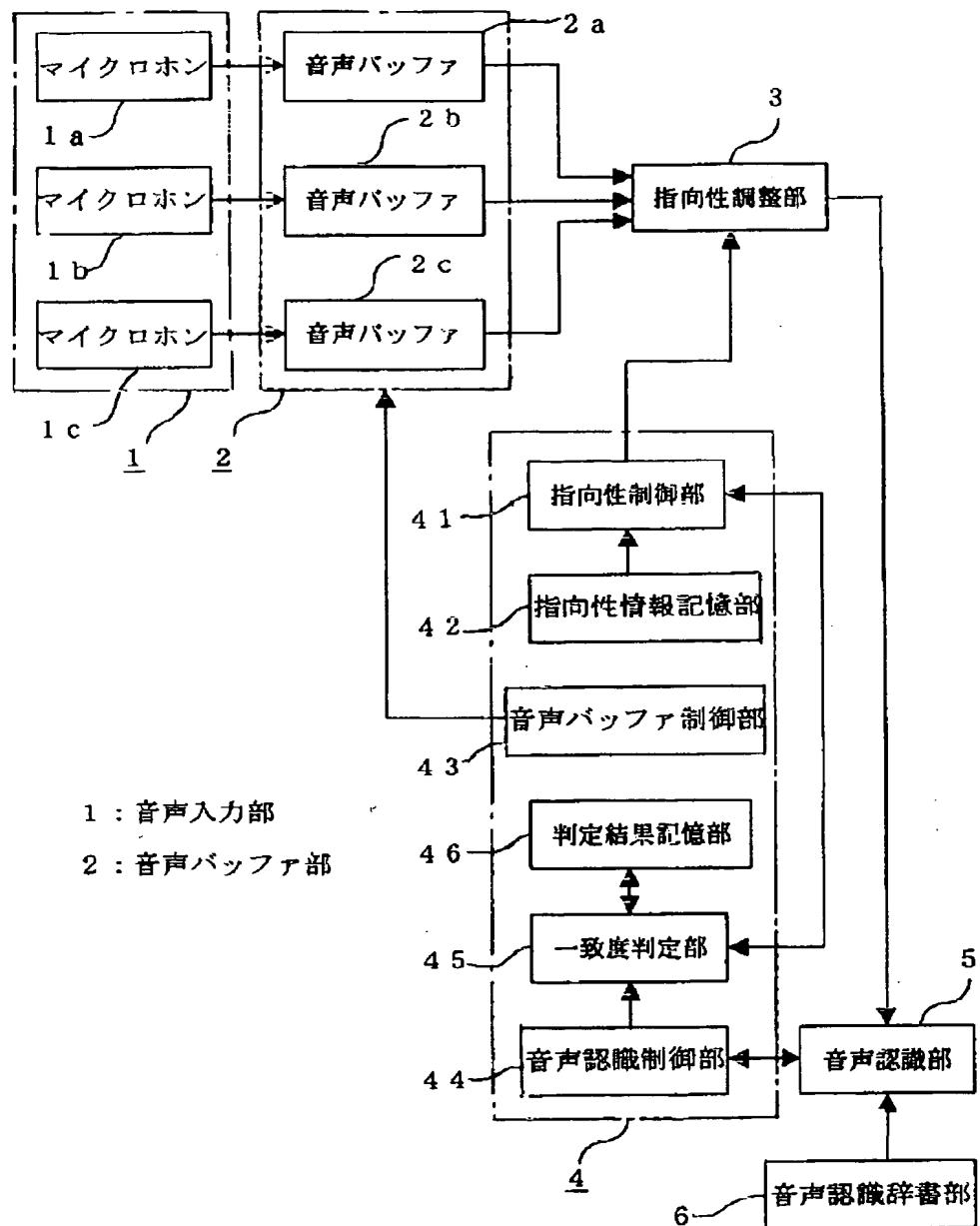
【符号の説明】

1 音声入力部、2 音声バッファ部、3 指向性調整部、5 音声認識部、6 音声認識辞書部、31 指向性利得調整部、41 指向性制御部、42 指向性情報記憶部、43 音声バッファ制御部、44 音声認識制御部、45 一致度判定部、46 判定結果記憶部、47 指向性利得制御部、48 指向性利得情報記憶部。

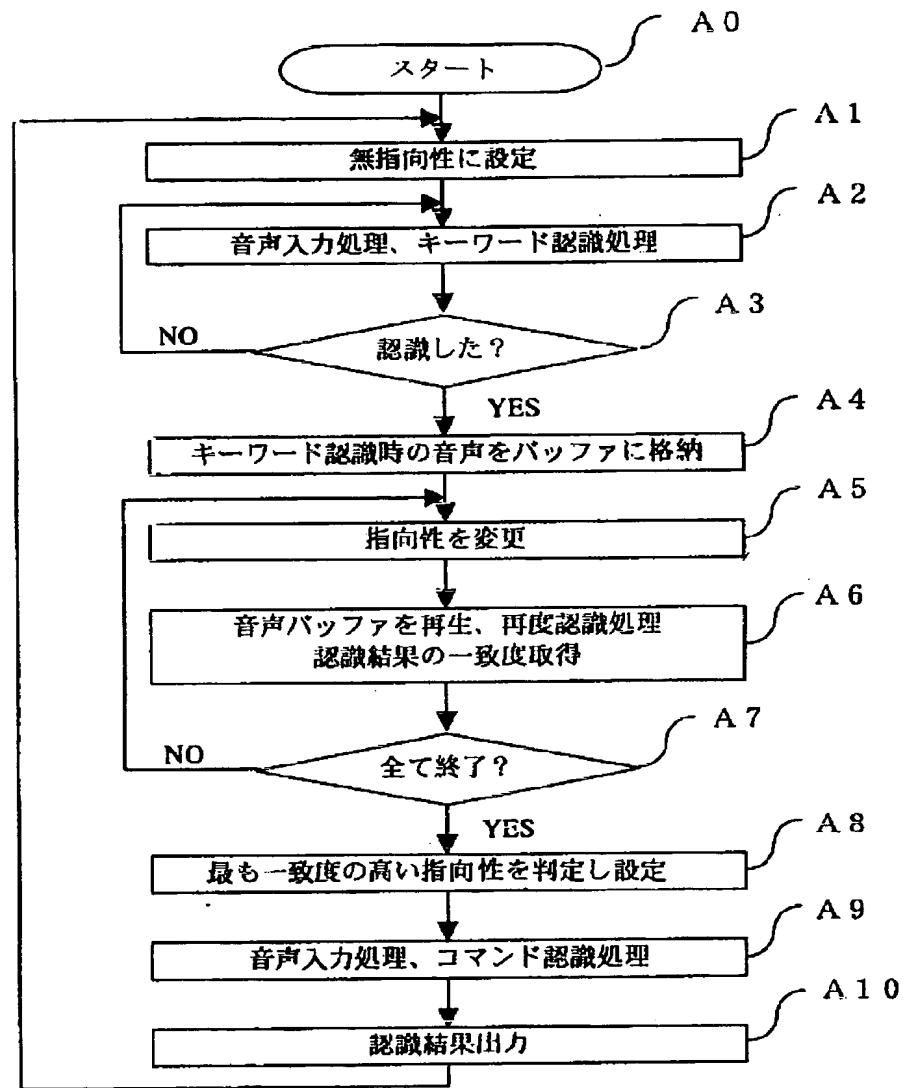
【図2】



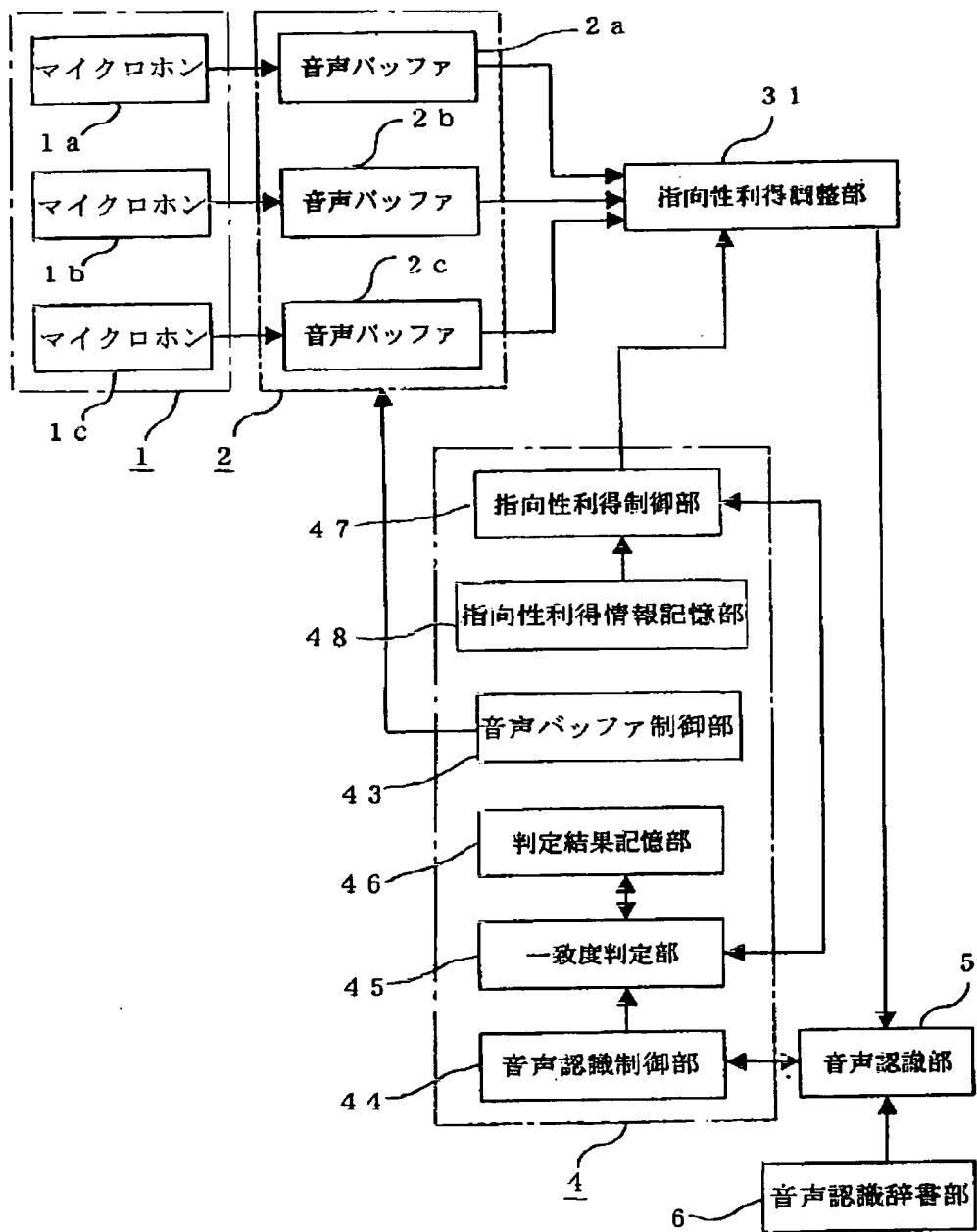
【図1】



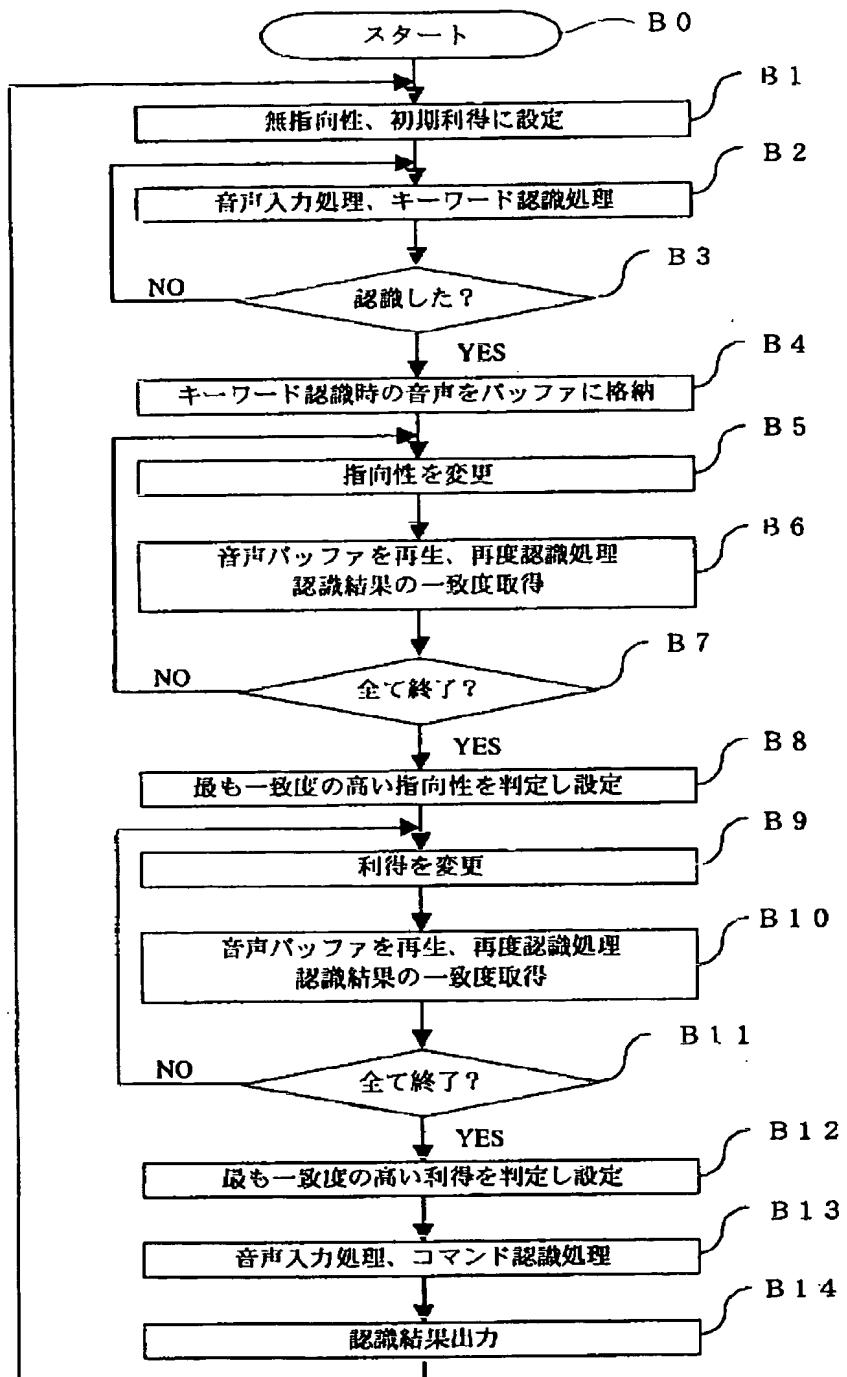
【図3】



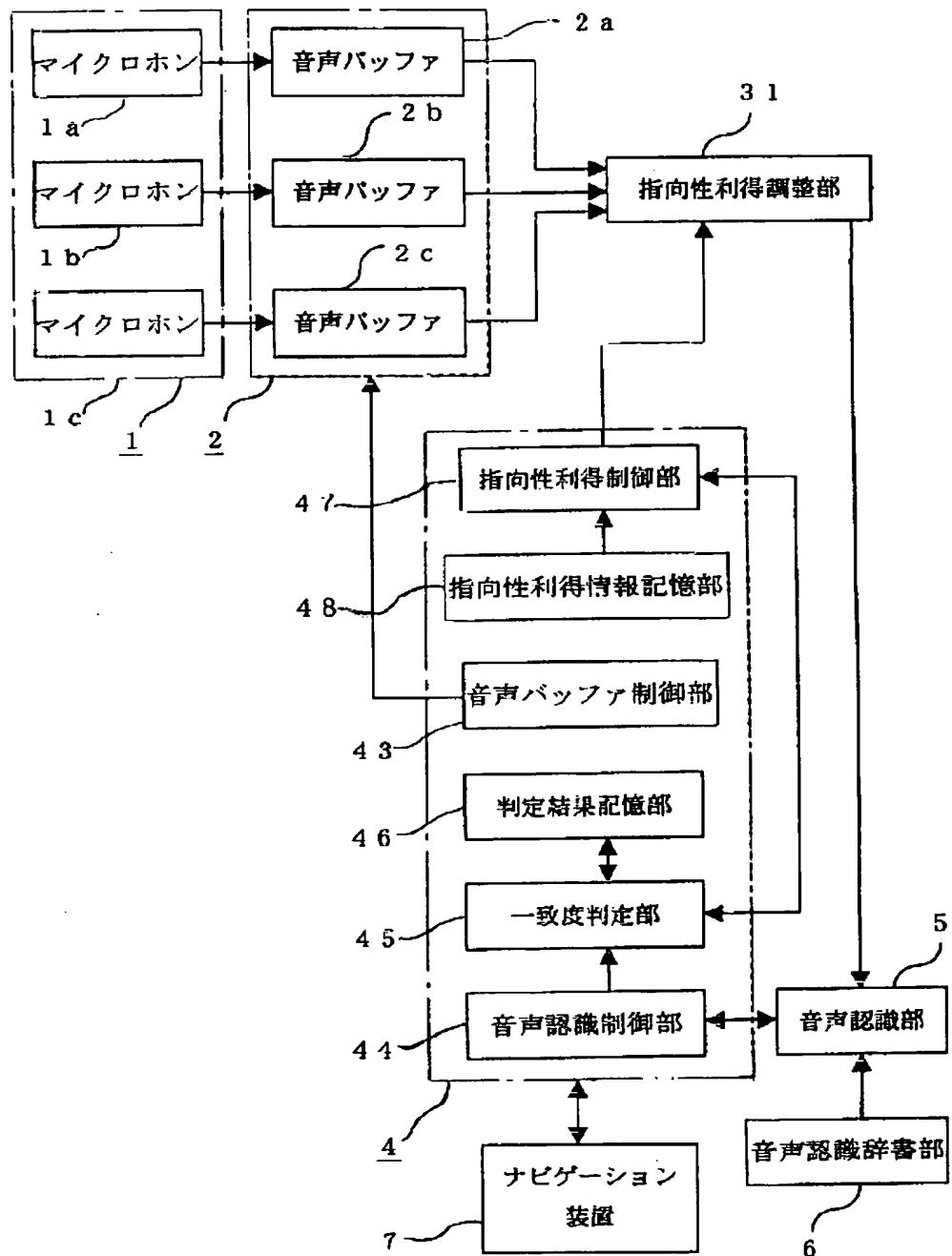
【図4】



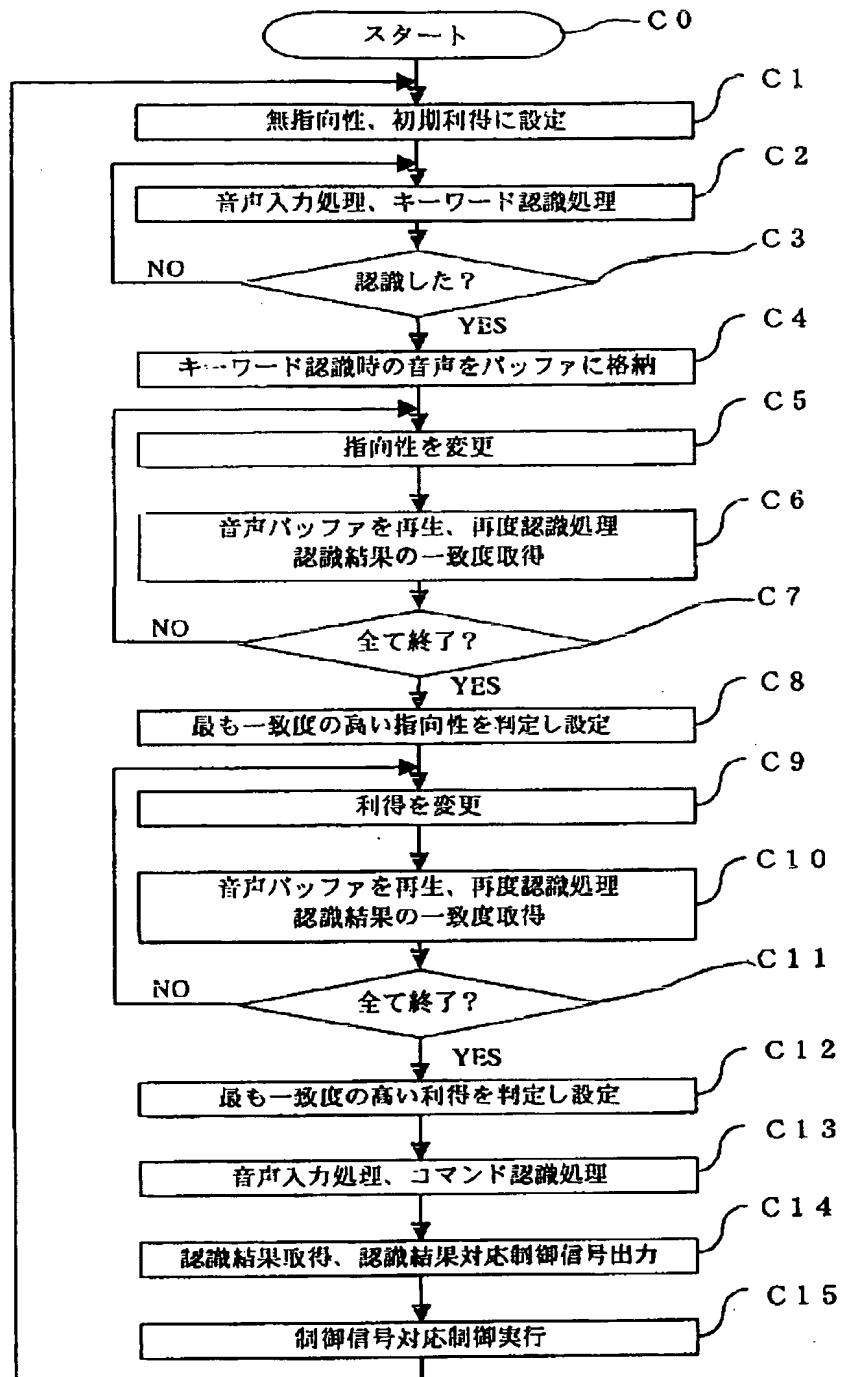
【図5】



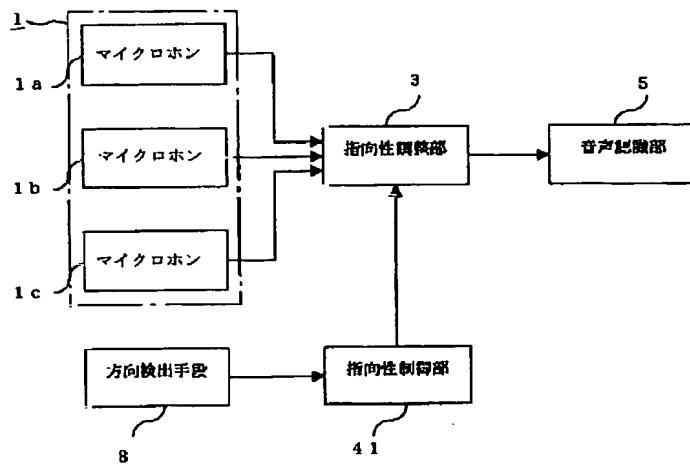
【図6】



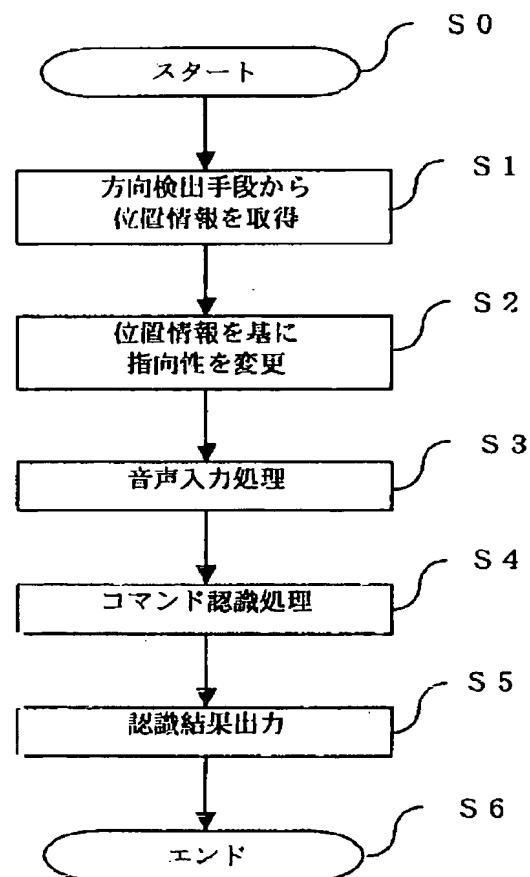
【図7】



【図8】



【図9】



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**